Probability and Independence ©

Terri Bittner, Ph.D.

The concept of *independence* is often confusing for students. This brief paper will cover the basics, and will explain the difference between *independent events* and *mutually exclusive events.*

Consider two events A and B. Once the events are defined, it is possible to talk about $P(A)$, $P(B)$, $P(A \cup B)$, $P(A \cap B)$, and so on. Let's recall that we can also represent the events themselves by Venn Diagrams. Also remember that

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

However, if A and B are *mutually exclusive*, it means that there is no *overlap* between A and B. This means that $A \cap B = \varnothing$ (the null set), and $P(A \cap B) = 0$. So, if A and B are mutually, exclusive, this implies that

$$P(A \cup B) = P(A) + P(B)$$

In other words, the probability that A or B happens is just the sum of the individual probabilities, but only if A and B are mutually exclusive.

Students often confuse the concept of mutually exclusive events with *independent* events. In fact, these are two different things. While mutual exclusivity of events means there is no overlap between the sets A and B, *independence* means that knowledge about what happened with event A does not affect event B, and knowledge about what happened with event *B* does not affect event *A*. This can be written mathematically as follows:

Theorem regarding independence of two events:

Two events *A* and *B* are *independent* if and only if $P(A \cap B) = P(A) \cdot P(B)$.

This means that we can conclude that $P(A \cap B) = P(A) \cdot P(B)$ if we are told that the events are independent (this is the *if* part of the theorem), and we can conclude that the events are independent if we know that $P(A \cap B) = P(A) \cdot P(B)$ (this is the *only if* part of the theorem).

The independence theorem is a very useful statement. We can also think of it as follows:

Independence of *A* and *B* implies that $P(A|B) = P(A)$, and $P(B|A) = P(B)$. This is just a mathematical way of saying that event *B* gives us no information about *A* and vice versa. While these conditional probability statements may be more satisfying intuitively that the formal theorem, the formal theorem is the more useful of the independence statements, since $P(A \cap B) = P(A) \cdot P(B)$ allows us to break up a complex statement ($P(A \cap B)$) into a relationship between two simpler statements ($P(A)$ and $P(B)$).

You may have heard about the idea of sampling with and without replacement. In general, sampling without replacement means that the samples are *not* independent. Sampling with replacement *may* mean that the samples are independent, but this is not necessarily the case in every problem. One other thing to note is that there is a situation where sampling without replacement can still result in independent samples, but *only if* the population (universe) you are selecting from is very, very large. In fact, this is only the case if the population is so large that taking one sample out of it does not materially affect the probability of your next draw, even if you don't "replace" after sampling.

These concepts are best illustrated via examples.

The tossing of a coin, the rolling of dice, and drawing marbles from an urn *with replacement* are all examples of independent events.

Suppose that you toss a fair coin 20 times in succession. Suppose that you toss "heads" the first 19 times. Interestingly, the probability of tossing "heads" on the twentieth toss is still ½. You can think of it this way: The coin cannot "remember" what was tossed the previous nineteen times, and so the previous tosses have nothing to do with the probabilities for the next toss. The same holds true for rolling dice.

Now, suppose you have 20 marbles in an urn. Suppose eight are red, ten are white, and two are blue. Define event A as drawing a blue marble on a single draw. B = the event of drawing a red marble, and C = drawing a white marble. Assuming that we replace the marbles after we draw them, then $P(A)$, $P(B)$, and $P(C)$ will always be the same no matter how many times we draw from the urn. Replacing the marbles means that the probabilities don't change from draw to draw, no matter what color we picked on the last draw. The urn is always "full" before every draw. However, if we start drawing marbles *without* replacement, then everything changes. Under this assumption, the events A, B, and C are *not* independent. Drawing a blue marble the first time (and not replacing it) changes the probabilities for drawing a red, white, or blue marble on the second draw, and so on. In other words, each draw has an effect on the next draw, and so the individual draws are not independent.

Example. Find the probability of drawing a red marble, then a white marble, then a blue marble, both with and without replacement.

Solution. Suppose we draw marbles with replacement. Then

$$P(A)=\frac{2}{20}=\frac{1}{10}$$
$$P(B)=\frac{8}{20}=\frac{2}{5}$$
$$P(C)=\frac{10}{20}=\frac{1}{2}$$

Note that $P(A)+P(B)+P(C)=\frac{1}{10}+\frac{2}{5}+\frac{1}{2}=1$, and this makes sense since we will always draw either a red, white, or blue marble. Also, it doesn't matter what order we draw our marbles in, since the probabilities don't change when we replace the marble after selection. Therefore,

$$\begin{aligned} P(\text{red, white, blue}) &= P(\text{red})\cdot P(\text{white})\cdot P(\text{blue}) \\ &= \frac{2}{5}\cdot\frac{1}{2}\cdot\frac{1}{10} \\ &= \frac{1}{50} \end{aligned}$$

(Notice the relationship between this example and our theorem about independece?)

The probability of a red, white, then blue marble $P(B, C, A)$ will be the same as the probability of white, then red, then blue, $P(C, B, A)$, which is the same as the probability of a blue, then red, then white, $P(A, B, C)$, etc. The order in which we draw the colors doesn't matter. The probabilities will all be the same under independence.

Now let's assume that we don't replace the marbles after each draw. Things change dramatically now, and we no longer have independent events. In fact, the events are dependent on each other because the probabilities for the next draw will change each time you draw a marble, and they'll be different depending on which color you draw each time. In this case,

$$P(\text{red, white, blue}) = \frac{2}{5} \cdot \frac{10}{19} \cdot \frac{2}{18} = \frac{4}{171}$$

The probability of drawing the red marble first is still $\frac{8}{20} = \frac{2}{5}$, since there are eight red marbles, and 20 marbles altogether. However, once we have drawn a red marble from the urn, there are now only 19 marbles left. Since ten of those 19 remaining marbles are white, the probability of drawing a white marble on the second draw is $\frac{10}{19}$, and the probability of drawing a blue marble on the third draw is similarly $\frac{2}{18} = \frac{1}{9}$.

It is easy to see that the answers can change quite a lot depending on whether samples are independent or not, so it is important to think carefully about the concept of independence before attempting problems involving sampling. Quite simply, ask yourself if the outcome of the first event will have any effect on the next event (or sample). If so, the events are dependent, and you cannot use the independence theorem. If the first event has no effect on the outcome of the next, as in the case of tossing a coin or rolling a pair of dice, then the events are *independent.* Never assume events are independent unless you are told they are, or unless independence is clear because of the nature of the events (as in the case of tossing a fair coin or rolling dice).